

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-185720

(43)Date of publication of application : 15.07.1997

(51)Int.Cl.

G06T 9/20

G06T 7/00

H04N 5/265

BEST AVAILABLE COPY

(21)Application number : 07-343894

(71)Applicant : CANON INC

(22)Date of filing : 28.12.1995

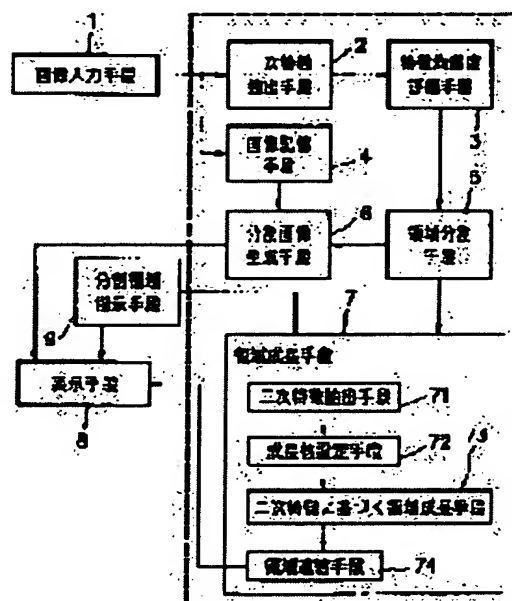
(72)Inventor : MATSUGI MASAKAZU

## (54) PICTURE EXTRACTION DEVICE

## (57)Abstract:

PROBLEM TO BE SOLVED: To easily and automatically segment a desired object from an inputted picture.

SOLUTION: A first feature extraction means 2 extracts the motion vector of the inputted picture, and a feature homogeneity evaluation means 3 and an area dividing means 5 set the distribution of the dimension and the direction of the motion vector to be an evaluation value to divide the area of the evaluation value within a prescribed limit. An area growing means 7 extracts the area of a color specified as a second feature from this divided area and connects the area, and a display means 8 displays it. A divided picture means 6 displays the divided area as a label picture. Thereby the object is simply and automatically segmented only by roughly specifying the position of the desired area based on the label picture by means of a mouse, etc. on a displayed picture.



## LEGAL STATUS

[Date of request for examination] 27.11.2002

[Date of sending the examiner's decision of rejection] 11.04.2006

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

**THIS PAGE BLANK (USPTO)**

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]

[Claim 1] The image extractor had a primary feature-extraction means extract the primary description of input image data, a field division means divide the field which has the above-mentioned primary description, a field selection means choose the field which fulfills predetermined conditions from the division field by which division was carried out [ above-mentioned ], a field growth means grow up the field by using as a nucleus the field by which selection was made [ above-mentioned ], and an image-display means display the image after the above-mentioned field growth.

[Claim 2] It is the image extractor according to claim 1 which carries out [ that the above-mentioned field growth means consists of a secondary feature-extraction means extract the secondary description from the field by which selection was made / above-mentioned /, a growth nucleus setting means set up the nucleus for field growth based on the above-mentioned secondary description, a secondary description field growth means to by which the secondary description of a proper performs field growth in the above-mentioned nucleus, and a connection means connect two or more fields after field growth, and ] as the description.

[Claim 3] the above-mentioned field division means -- distribution of the above-mentioned primary description data -- abbreviation -- the image extractor according to claim 1 characterized by performing division which carries out grouping of the homogeneous field.

[Claim 4] The above-mentioned field division means is an image extractor according to claim 1 characterized by making the field more than predetermined size into the above-mentioned division field among the connection fields obtained by carrying out connection processing of the field which has the above-mentioned primary description.

[Claim 5] It is the image extractor according to claim 1 which the above-mentioned field division means has the means which carries out label attachment of each of the above-mentioned division field, establishes an image generation means to generate the label image data which has the predetermined description, all over the same field of the above-mentioned label, and is characterized by for the above-mentioned image display means to display the above-mentioned label image data and the above-mentioned input image data in piles.

[Claim 6] It is the image extractor according to claim 1 which the above-mentioned field selection means chooses what has the maximum area among the above-mentioned division fields, and is characterized by the above-mentioned image display means displaying the image data of the field by which selection was made [ above-mentioned ].

[Claim 7] It is the image extractor according to claim 1 which the above-mentioned field selection means chooses what has average brightness or an average hue in the predetermined range among the above-mentioned division fields, and is characterized by the above-mentioned image display means displaying the image data of the field by which selection was made [ above-mentioned ].

[Claim 8] It is the image extractor according to claim 1 which the above-mentioned input image data is time series image data, and is characterized by the above-mentioned primary description being motion vector distribution.

[Claim 9] It is the image extractor according to claim 1 which the above-mentioned input image

**THIS PAGE BLANK (USPTO)**

data is time series image data, and is characterized by the above-mentioned primary description being subtraction-image distribution of input image data.

[Claim 10] The above-mentioned input image data is an image extractor according to claim 1 which consists of two or more image data from which the view location in the same scene differs, and is characterized by the above-mentioned primary description being parallax distribution between two or more above-mentioned image data.

[Claim 11] The above-mentioned field selection means is the image extractor according to claim 5 carry out having a marker generation means for displaying a marker on the location which directed with the above-mentioned field selection directions means based on the above-mentioned label image displayed on the field selection directions means and the above-mentioned image-display means on the above-mentioned image-display means, and a judgment means judge a field including the above-mentioned marker position to be a selection field as the description.

[Claim 12] The above-mentioned input image data is an image extractor according to claim 1 obtained from an image pick-up means.

[Claim 13] The above-mentioned image pick-up means is an image extractor according to claim 12 which is what is prepared in the mobile.

[Claim 14] The above-mentioned image pick-up means is an image extractor according to claim 12 which is what has a compound eye image pick-up system.

[Claim 15] The image extractor according to claim 14 whose above-mentioned primary description is parallax distribution of the corresponding points between each image pick-up system in the above-mentioned compound eye image pick-up system.

---

[Translation done.]

THIS PAGE BLANK (USPTO)

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the image extractor which performs logging of an image and the extract of a border line.

[0002]

[Description of the Prior Art] Conventionally, the approach of placing the object as an approach of starting the field of a desired object into the background which has a fixed color (blue), and carrying out chroma-key processing from an image, is learned well. Moreover, the range started along with the border line of an object is specified by the help, and the methods (JP,6-251149,A, JP,7-37086,A, etc.) of asking a key for information, such as a color component and brightness, for the exact location of logging by the image processing etc. are learned. As an example which attained a certain amount of automation, an operator specifies one point and the profile trace direction of [ on a border line ] indicated by JP,7-36173,A, and there is the approach of tracking a profile automatically based on the inclination data of shade distribution of the image obtained by differential processing. This is the image logging approach which interrupted processing, inputs the trace direction etc. anew based on directions of an operator when a profile branches, or when ambiguous, and was equipped with motility and interactive nature.

[0003] Moreover, in image-processing tools, such as Photoshop (Adobe company), if one in a field to extract is specified with a mouse in the case of logging from a background, including the point, a color component can connect the field in predetermined within the limits automatically, and the field of different color component range can be similarly formed and merged. As an approach of further others, the background image is memorized and there are methods (JP,4-216181,A, JP,6-266841,A, JP,6-268902,A, etc.) of extracting a quiescence body image from a subtraction image with the present image.

[0004]

[Problem(s) to be Solved by the Invention] However, the approach by the chroma-key processing mentioned above had problems, like the part is missing at the time of an image extract, when it had that a background must be a specific color and the color as a background with the same object. Moreover, the method of specifying a border line by the operator essentially extracts most profile configurations of an outline by the help, and the extract from the image accompanied by the body of a complicated configuration, a motion, or deformation was difficult. Moreover, the approach of pursuing this border line had the indispensable management by the operator, when an object had a complicated pattern, and when it had further a background and texture (pattern) structure to which the object was similar and processing by the help took place frequently, it had the problem that the semantics of automation was lost.

[0005] By the approach of carrying out automatic extracting of the field where the color component approached using the image-processing tool, there was a problem that it was generally impossible to cut down an object image by one actuation, and the field merge covering many times was needed. Furthermore, it was difficult stability and to not necessarily start only a target object correctly according to factors, such as a noise, by the approach using difference with a background image. the image obtained while performing camera actuation of panning etc.

THIS PAGE BLANK (USPTO)



when especially a camera was used as an image input means -- setting -- difference -- it could not respond only by processing but there was a problem that increase of the amount of operations -- global matching processing is needed -- was not avoided.

[0006] Then, the 1st purpose of this invention is only directing the location of an object roughly and making automatic extracting of an image possible, without pursuing the border line by the help or the image processing. The 2nd purpose of this invention is performing efficiently narrowing down of an object image, and the minute extract from a background by using the classification of two or more description data.

[0007] the field where the 3rd purpose of this invention was divided based on the field which had complete set of description data roughly -- as the nucleus candidate field of a field grown method -- directions -- it is supposing that it is selectable and making automatic extracting of a specific image possible. the field divided based on little data, without the 4th purpose of this invention evaluating whenever [ homogeneity / of the description ] -- as the nucleus candidate field of a field grown method -- directions -- it is supposing that it is selectable and making automatic extracting of a specific image possible.

[0008] The 5th purpose of this invention is making easy the selection directions for [ by the operator ] an extract by displaying in piles the field divided roughly on an input image. The 6th purpose of this invention is it carrying out automatic extracting, using as the candidate for an extract the maximum area field which has the same image attribute, and realizing simplification and improvement in the speed of processing. The 7th purpose of this invention is an average brightness and color component's using the image field in the predetermined range as a field growth nucleus, carrying out automatic extracting, and realizing the simplification and improvement in the speed of processing. The 8th purpose of this invention is realizing the image extract based on the 1st purpose at the time of extracting a subject of operation from a time series image.

[0009] The 9th purpose of this invention is realizing the image extract based on the 1st purpose at the time of extracting a subject of operation from a time series image at a high speed. The 10th purpose of this invention is realizing the image extract based on the 1st purpose, when the average distance of the for a background and for an extract from a view location differs enough. The 11th purpose of this invention is realizing the image extract based on the 1st purpose by rough directions selection interactively.

[0010]

[Means for Solving the Problem] A primary feature-extraction means extract the primary description of input image data in this invention, a field division means divide the field which has the above-mentioned primary description, a field selection means choose the field which fulfills predetermined conditions from the division field by which division was carried out [ above-mentioned ], a field growth means grow up the field by using as a nucleus the field by which selection was made [ above-mentioned ], and an image-display means display the image after the above-mentioned field growth have prepared.

[0011]

[Function] According to this invention, the field which has the primary description first from input image data is divided, and the field which fulfills further predetermined conditions among the divided field is chosen, and, for example based on the secondary description, expansion growth of the field is further carried out from the selected field, it goes, and the image after growth is displayed. Therefore, only by carrying out tab control specification of the field which should be extracted on the displayed image roughly, an object image can be cut down automatically easily (extract) and can be carried out.

[0012]

[Embodiment of the Invention] Drawing 1 is drawing showing the example of a system configuration of this invention as a gestalt of the 1st operation notionally. In drawing, 1 is an image input means and an image pick-up means or an image database is used. For the field division means based on [ 2 / 3 / the primary description data extraction means in input image data, and / whenever / homogeneity / of the description data / based on the image storage means of an input image in an evaluation means and 4 ] whenever [ description homogeneity ] in

THIS PAGE BLANK (USPTO)

5, and 6, as for a field growth means and 8, a division image generation means and 7 are [ a display means and 9 ] the selection directions means of a division field, and pointing means, such as a mouse, are used.

[0013] In addition, field growth is evaluating whenever [ homogeneity / of the characteristic quantity in the predetermined field in an image ], and performing division of a field, and a merge. While the description of the gestalt of this operation extracts roughly the image field which has extract possibility by extracting the primary description data in an input image at a high speed, and performing field division of an image based on the distribution It is performing field growth based on the secondary characteristic quantity (characteristic quantity from which the primary description, classification, an attribute, etc. differ) which uses a candidate field as a nucleus as a following step for a more minute extract. In the field extracted based on the primary description data, image information required to perform field growth based on the secondary description shall be contained.

[0014] In case the primary description data extraction means 2 extracts the field used as the nucleus at the time of performing processing by field growth means 7 to explain later at a high speed and actually extracts the image field for specification, the rough selection and the directions by the operator can perform it. Here, it explains as what extracts the motion vector in screen top each point as an example of the primary description data.

[0015] The distribution about the magnitude and the direction of a motion vector of [ in a predetermined field ] is used for the evaluation means 3 as an evaluation value whenever [ description homogeneity ]. Compaction and the increase in efficiency of the processing time are attained by setting up the size of the field (block) to evaluate appropriately based on the target size in a screen. That is, as compared with the target size in every direction [ in a screen ], about several [ 1/ ] to 1/10 division field extent is used typically. You may make it set up this value suitably beforehand by the operator side.

[0016] If the field division means 5 has the evaluation value of whenever [ above-mentioned homogeneity ] smaller than a predetermined threshold, it will consider that the field is homogeneous and will connect the fields which have an evaluation value in predetermined tolerance whenever [ homogeneity ]. Moreover, the field outside tolerance forms the lump (plurality is good) of a field which can dissociate and can represent with one motion vector. The input image picturized by drawing 7 (a) with the fixed camera is shown, and the result of the field division performed based on the motion vector is shown in this drawing (b). The result of image logging by the field growth based on the secondary descriptions, such as a color component, is shown in this drawing (c).

[0017] In addition, the field after division may not necessarily reflect the configuration of an actual object in a precision. Therefore, although the primary feature-extraction means 2 may be somewhat inferior to the precision (resolution) of the secondary feature-extraction means 71 used later, to be able to extract at a high speed is desired. However, in case it is that some background images are not intermingled in a division field, and field growth, it is desirable to include all required image attributes (secondary description). For the former, it is good after field division also as a division field used for the processing which follows what reduced or deleted each field at a fixed rate.

[0018] The division image generation means 6 gives the image data (for example, a different color, a different hatching pattern, etc.) of an attribute which is different corresponding to two or more fields divided based on the primary description data, respectively, and it is displayed that it laps with input image data on the display means 8. Label attachment is performed to the specifically divided field, and it processes smearing away the inside of a mask by the fixed pattern etc. by using the same label field as mask data. By doing in this way, an operator can distinguish easily the candidate field which should be directed by viewing, and assignment for [ by the division field directions means 9 ] an extract becomes certain (since it is beforehand classified with the same label), and easy. Although a mouse etc. is typically used as a division field directions means 9, especially when using a touch panel mold display, there is no need.

[0019] The field growth means 7 is divided based on the primary description data as mentioned above. The secondary description (attribute) from the selected field It consists of a growth

THIS PAGE BLANK (USP: 0,

nucleus setting means 72 in the case of performing field growth based on a secondary feature-extraction means 71 to extract, and its secondary description, a field growth means 73 to perform field growth for the secondary description of a proper based on each growth nucleus, and a field connection means 74 to connect each field after growth. In addition, field growth means performing division of a field, and a merge according to whenever [ homogeneity / of the description data ], and, on the other hand, merge processing is not necessarily given to a target. Moreover, although there are some which have the thing which has geometric structure usually, for example or luminance distribution, and the high order description which processes a color component and is acquired that the secondary description should just be a fundamentally different thing from the primary description, it is not necessarily limited to these, for example, color component information etc. is used suitably. However, it is satisfactory, if the minute image division by field growth is possible even if it is the same classification (here motion vector) as the primary description (refer to the gestalt of the 5th operation).

[0020] After extracting the secondary description as mentioned above, the initial field of a proper is set as each with the growth nucleus setting means 72. When secondary characterized by color component information, the secondary description plurality should grow up to be may exist in the extracted candidate field. For example, in the case of a person's image, it has a hue which is different on the average according to parts, such as a face, the hair, and clothes, respectively, but it considers as the profile part which has the secondary description from which the candidate field which was extracted as a growth nucleus in that case based on the primary description data, and was chosen with the division field directions means 9 differs, or its near field. After equalizing within the near field block of predetermined size beforehand in the case of field growth, you may make it grow up. The effect of fluctuation of the image data (attribute) in the resistance and pixel size order to a noise can be eased by this, and extract precision can be raised.

[0021] Drawing 2 shows the gestalt of the 2nd operation and shows the means 2-9 which materialized other notional means 2-9 of drawing 1 at the time of using the image pick-up means 1 for the image input means 1 of drawing 1. It is the case where made the primary description data into the motion vector here, and the secondary description data are made into a hue. In addition, about the detection algorithm of a motion vector, although explanation is omitted since it is not the chief aim of this invention, especially approaches with rapidity as shown in a "robot vision" (the Yanaida work, Shokodo issue), such as a gradient method and a correlation technique, are desired, for example.

[0022] It may dissociate from the background for specification based on the difference of a motion vector in case all that include actuation, the separation from the quiescence background of the object accompanied by a motion, or a background as an application example of the gestalt of this operation move. for example, -- the case where a certain scene is picturized while the image pick-up means 1 is installed in the body, a robot, or a car body and the image pick-up means 1 moves -- general -- the screen whole -- crossing -- un--- zero -- a motion vector exists. In such a case, if a background and the distance from the image pick-up means 1 of an object are especially separated enough, it can move like the above-mentioned and field division based on a vector can be performed comparatively with high precision. When there are few objects which carry out the object which moves especially, or a different motion from a perimeter, the gestalt of this operation is effective.

[0023] Drawing 3 shows the gestalt of the 3rd operation. With the gestalt of this operation, parallax distribution of the corresponding points between right-and-left camera images is used as primary description data, using a compound eye image pick-up system (stereoscopic camera) as an image input means. Stereo image mensuration is shown in "three dimensional image measurement" (Inokuchi, the Sato work, Shokodo) etc. Moreover, he is trying for field area with the homogeneous parallax of the photographic subject which should be extracted to serve as max by establishing the image pick-up device control means 10, and controlling the angle of convergence of a compound eye camera, a focal distance, the base length, etc. by the gestalt of this operation based on the parallax data of the parallax extract means 2. While an operator observes the label image on the compound eye image display means 8, it may be made to

THIS PAGE BLANK (US:10)

perform this control. Since parallax is acquired only about the field which overlapped between right-and-left images, the display of the candidate field after dividing into the field where a parallax value is homogeneous is performed only in a duplication field. Field growth is performed in the image selection section 11 by choosing one candidate for an extract among right-and-left images. Moreover, although the hue is used about the secondary description here, it is not limited to this.

[0024] Drawing 4 shows the gestalt of the 4th operation. By obtaining a stereo image as time series image data, concomitant use with the gestalt of the 2nd operation is possible for the gestalt of implementation of the above 3rd. Drawing 4 shows the example of a system configuration. Here, the primary description data extraction means 2 extracts a motion vector and parallax distribution, and each description data performs clustering based on whenever [ homogeneity ] with field division and the label attachment means 5a and 5b. An input image is shown in drawing 8 (a), and the field division result based on the primary description (parallax and motion vector) is shown in this drawing (b).

[0025] Next, the label image by which the deflection between representation characteristic quantity (parallax, motion vector) was computed, field division was carried out based on the description data with larger deflection, and the area size after division was generated about the cluster beyond a predetermined value with the deflection calculation means 12a and 12b is used in the case of the candidate field selection for an extract. the distance with the body almost same [ this ] from the compound eye animation image pick-up means 1 in a screen -- being distributed (the deflection of parallax being smallness), when the photographic subject which should be extracted has a remarkable motion to a perimeter the direction of the field division using a motion vector -- reliability -- high -- moreover, distance with a perimeter and a background -- the distance of the compound eye animation image pick-up means 1 and an object -- comparing -- enough -- being large (the deflection of parallax being size), when all do not show a remarkable motion relatively (the deflection of a motion vector is smallness) It is because the field division based on parallax is desirable.

[0026] Drawing 8 (c) shows a processing result, chooses the largest object of a parallax value here, and shows the result of having performed field growth using the secondary description minute amount (color component). Thereby, the image data which was suitable for the image extract according to the size of the existence of a motion of an object and dispersion of distance distribution can be used properly. Moreover, you choose the classification of the description data beforehand used based on the primary description data, and may also narrow down image pick-up modes (a still picture and a stereo image pick-up, or an animation, a monocular vision image pick-up, etc.). Moreover, in using only the dynamic-image data from one camera, it applies processing with the gestalt of the 2nd operation. In addition, as an image input means, it may not be limited to a compound eye image pick-up system, but the image with parallax based on a three-dimensions geometric model generated by computer graphics is sufficient.

[0027] Drawing 5 shows the gestalt of the 5th operation. The gestalt of this operation uses an input image as the time series image from the image pick-up means 1, such as an ocellus camera by which fixed installation was carried out. If it considers as the primary description data, the outline field (a location and magnitude) of a part with the primary description data lost motion is extracted using inter-frame subtraction-image data. An image is extracted with the application of the field grown method based on the selection by label attachment, the display means 8, mouse, or the BOINTINGU means 9 of a candidate field, and the secondary description over a candidate field.

[0028] drawing 5 -- setting -- difference -- the primary description data obtained by making binary the subtraction image obtained by data extraction means 2a with a predetermined threshold with binary-ized means 2b carry out connection processings of them, such as etc., the distance between segments is below the number of predetermined pixels -- in accordance with fixed criteria, without expressing coarse division of a motion vector to a quiescence background, therefore evaluating whenever [ homogeneity ]. It carries out label attachment, using as a candidate field the field across which the description data (non-zero data) existed the front and

THIS PAGE BLANK (USPTO)



behind objective, and carried out grouping, respectively, and they faced, and it leaves as a division field. [ which were typically moved in the predetermined direction ] Furthermore, the field below predetermined size is omitted. Although the hue is used like the gestalt of each operation as secondary description data, it is not limited to this.

[0029] Drawing 6 shows the gestalt of the 6th operation. the gestalt of this operation -- an input image -- receiving -- the primary description data (a motion vector --) of multiplex resolving and a hierarchical type Extract processings, such as parallax, are performed and it is low resolution data (for example, after forming an image into low resolving by infanticide processing or equalization) as a primary description. The secondary description uses the high resolution data of the same classification, such as computing the motion vector for every block by detecting a motion vector or dividing image data into the large block of size. According to the gestalt of this operation, the high-speed processing by one feature-extraction processing algorithm is possible.

[0030] Drawing 6 shows the system configuration at the time of using a motion vector as primary description data. Minute logging is performed by performing a setup (growth nucleus setting means 72) of a field growth nucleus with a low resolution from evaluation (whenever [ homogeneity ] evaluation means 3) of whenever [ motion vector homogeneity ], and performing field growth (field growth means 73) with high resolution. Like the gestalt of the 2nd operation, it cannot be overemphasized that field growth means the merge or separation processing of subregion based on whenever [ homogeneity ]. Since especially growth nuclei are gestalt \*\*\*\*\* of this operation, and a field fundamentally represented with the same motion vector, as compared with the field growth using a color component, a growth nucleus setup is simplified very much, and more nearly high-speed processing is possible for them also in field growth processing.

[0031]

[Effect of the Invention] As explained above, according to the 1st invention, the image for specification in the background of arbitration can be automatically extracted only by directing a rough location. Moreover, according to the 2nd invention, an object image can be efficiently extracted by utilizing the feature parameter of classification with which plurality differs. Moreover, according to the 3rd invention, when especially the description distribution data present some homogeneous fields and make them selectable, automatic extracting of the same image as the 1st invention can be performed. According to the 4th invention, the same image logging as the 2nd invention can be performed, without performing evaluation of whenever [ homogeneity ].

[0032] According to the 5th invention, an operator can perform easily assignment of the field which should be extracted. According to the 6th invention, a setup of the image candidate field which should be extracted is automatable. According to the 7th invention, narrowing down and a setup of an image candidate field which should be extracted are automatable.

[0033] According to the 8th invention, the extract and narrowing down of a candidate field based on the target motion vector component can be performed. According to the 9th invention, the object image candidate field from the field where the motion was detected can be extracted, and the same effectiveness as the 8th invention can be acquired. According to the 10th invention, narrowing down of an image candidate field and a setup are automatable from a stereo image with parallax. According to the 11th invention, the operability at the time of only rough directions performing the alternative extract of an object domain image can be raised.

---

[Translation done.]

THIS PAGE BLANK (USPTO)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-185720

(43) 公開日 平成9年(1997)7月15日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 T 9/20			G 0 6 F 15/70	3 3 5 Z
	7/00		H 0 4 N 5/265	
H 0 4 N 5/265			G 0 6 F 15/70	3 3 0 A

審査請求 未請求 請求項の数15 O L (全 8 頁)

(21) 出願番号 特願平7-343894

(22) 出願日 平成7年(1995)12月28日

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 真継 優和

東京都大田区下丸子3丁目30番2号 キヤ  
ノン株式会社内

(74) 代理人 弁理士 國分 孝悦

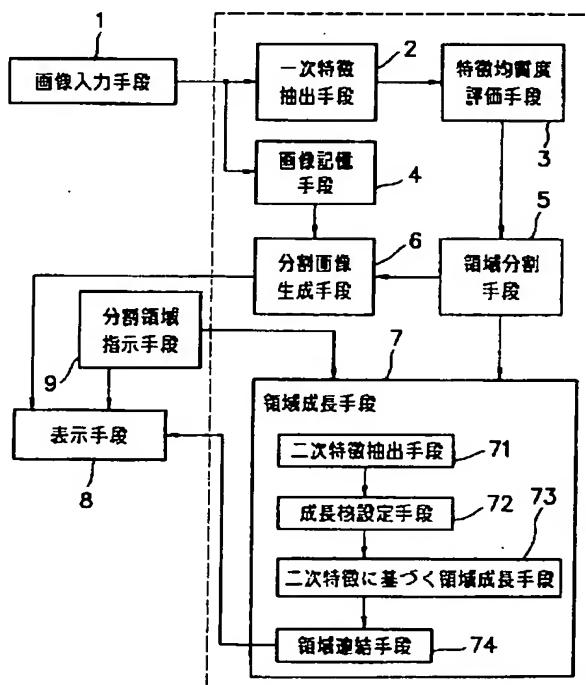
(54) 【発明の名称】 画像抽出装置

(57) 【要約】

【課題】 入力画像から所望の対象物を容易に自動的に切り出す。

【解決手段】 一次特徴抽出手段2は入力画像の動きベクトルを抽出し、特徴均質度評価手段3、領域分割手段5は動きベクトルの大きさと方向の分布を評価値とし、評価値が所定範囲内の領域を分割する。この分割領域から領域成長手段7により、二次特徴として指定された色の領域を抽出し、その領域を連結して表示手段8で表示する。分割画像手段6は上記分割領域をラベル画像として表示させる。

【効果】 表示画像上でマウス等によりラベル画像に基づいて所望の領域の位置を大まかに指定するだけで対象物を簡単に自動的に切り出すことができる。



**【特許請求の範囲】**

**【請求項 1】** 入力画像データの一次特徴を抽出する一次特徴抽出手段と、

上記一次特徴を有する領域を分割する領域分割手段と、  
上記分割された分割領域から所定条件を満たす領域を選択する領域選択手段と、

上記選択された領域を核としてその領域を成長させる領域成長手段と、

上記領域成長後の画像を表示する画像表示手段とを備えた画像抽出装置。

**【請求項 2】** 上記領域成長手段は、上記選択された領域から二次特徴を抽出する二次特徴抽出手段と、上記二次特徴に基づいて領域成長のための核を設定する成長核設定手段と、上記核に固有の二次特徴により領域成長を行う二次特徴領域成長手段と、領域成長後の複数の領域を連結する連結手段とからなることを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 3】** 上記領域分割手段は、上記一次特徴データの分布が略均質な領域をグループ化する分割を行うことを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 4】** 上記領域分割手段は、上記一次特徴を有する領域を連結処理して得られる連結領域のうち所定サイズ以上の領域を上記分割領域とすることを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 5】** 上記領域分割手段は、上記分割領域のそれぞれをラベル付けする手段を有し、上記ラベルの同じ領域中に所定の特徴を有するラベル画像データを生成する画像生成手段を設け、上記画像表示手段は、上記ラベル画像データと上記入力画像データとを重ねて表示することを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 6】** 上記領域選択手段は上記分割領域のうち最大面積を有するものを選択し、上記画像表示手段は上記選択された領域の画像データを表示することを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 7】** 上記領域選択手段は上記分割領域のうち平均的な輝度または色相が所定の範囲にあるものを選択し、上記画像表示手段は上記選択された領域の画像データを表示することを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 8】** 上記入力画像データは時系列画像データであり、上記一次特徴は動きベクトル分布であることを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 9】** 上記入力画像データは時系列画像データであり、上記一次特徴は入力画像データの差分画像分布であることを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 10】** 上記入力画像データは同一シーン中の視点位置の異なる複数の画像データから成り、上記一次特徴は上記複数の画像データ間の視差分布であることを特徴とする請求項 1 に記載の画像抽出装置。

**【請求項 11】** 上記領域選択手段は領域選択指示手段と、上記画像表示手段に表示された上記ラベル画像に基づいて上記領域選択指示手段で指示した位置にマーカを上記画像表示手段上に表示するためのマーカ生成手段と、上記マーカ位置を含む領域を選択領域と判定する判定手段とを有することを特徴とする請求項 5 に記載の画像抽出装置。

**【請求項 12】** 上記入力画像データは撮像手段から得られたものである請求項 1 に記載の画像抽出装置。

**【請求項 13】** 上記撮像手段は移動体に設けられているものである請求項 12 に記載の画像抽出装置。

**【請求項 14】** 上記撮像手段は複眼撮像系を有するものである請求項 12 に記載の画像抽出装置。

**【請求項 15】** 上記一次特徴が上記複眼撮像系における各撮像系間における対応点の視差分布である請求項 14 に記載の画像抽出装置。

**【発明の詳細な説明】****【0001】**

**【発明の属する技術分野】** 本発明は画像の切り出しや輪郭線の抽出を行う画像抽出装置に関する。

**【0002】**

**【従来の技術】** 従来、画像から所望の対象物の領域を切り出す方法として、その対象物を一定色（青色）を有する背景中に置いてクロマキー処理する方法がよく知られている。また、対象物の輪郭線に沿って切り出す範囲を手により指定し、切り出しの正確な位置を色成分、輝度などの情報を手がかりに画像処理で求める方法（特開平 6-251149 号公報、特開平 7-37086 号公報など）などが知られている。ある程度の自動化を達成した例としては、特開平 7-36173 号公報に開示されるようにオペレータが輪郭線上の一点と輪郭追跡方向とを指定し、微分処理によって得られる画像の濃淡分布の勾配データに基づいて輪郭を自動追跡する方法がある。これは輪郭が分岐する場合または曖昧な場合には処理を中断してオペレータの指示に基づいて改めて追跡方向等を入力するもので、自動性とインタラクティブ性とを備えた画像切り出し方法である。

**【0003】** また、Photoshop（Adobe 社）などの画像処理ツールにおいては、背景からの切り出しの際、抽出したい領域内の一点をマウスで指定すると、その点を含み色成分が所定範囲内にある領域を自動的に連結し、かつ異なる色成分範囲の領域も同様にして形成、併合することができる。さらに他の方法として、背景画像を記憶しておいて、現画像との差分画像から静止物体画像を抽出する方法（特開平 4-216181 号公報、特開平 6-266841 号公報、特開平 6-268902 号公報など）がある。

**【0004】**

**【発明が解決しようとする課題】** しかしながら、上述したクロマキー処理による方法は、背景が特定色でなけれ

ばならないことや、対象物が背景と同じ色を有する場合は、その部分が画像抽出時に欠落することなどの問題があった。また、オペレータによって輪郭線を指定する方法は、本質的に人手により概略の輪郭形状を殆ど抽出するものであり、複雑な形状の物体や、動きや変形を伴う画像からの抽出は困難であった。また、この輪郭線を追跡する方法は、対象が複雑な模様を有する場合には、オペレータによる対処が不可欠であり、さらに背景と対象が類似したテクスチャ（模様）構造を有したりする場合など人手による処理が頻繁に起こる場合は、自動化の意味が失われるという問題があった。

【0005】画像処理ツールを用いて色成分の近接した領域を自動抽出する方法では、一回の操作で対象画像を切り出すことは一般的に不可能であり、多数回にわたる領域併合が必要になるという問題があった。さらに、背景画像との差分を用いる方法では、ノイズなどの要因により、必ずしも目標とする対象だけを安定かつ正確に切り出すことが困難であった。特にカメラを画像入力手段として用いた場合、パンニングなどのカメラ操作を行いながら得た画像などにおいては、差分処理だけでは対応できず大域的なマッチング処理が必要となるなど演算量の増大が回避できないという問題があった。

【0006】そこで、本発明の第1の目的は、人手または画像処理による輪郭線の追跡を行わずに、大まかに対象物の位置を指示するだけで、画像の自動抽出を可能とすることである。本発明の第2の目的は、複数の特徴データの種別を用いることにより、対象物画像の絞り込みと背景からの精細抽出とを効率的に行うことである。

【0007】本発明の第3の目的は、特徴データが大まかに揃った領域に基づいて分割された領域を領域成長法の候補領域として指示選択可能とし、特定画像の自動抽出を可能とすることである。本発明の第4の目的は、特徴の均質度を評価せずに少ないデータに基づいて分割された領域を領域成長法の候補領域として指示選択可能とし、特定画像の自動抽出を可能とすることである。

【0008】本発明の第5の目的は、大まかに分割された領域を入力画像上に重ねて表示することにより、オペレータによる抽出対象の選択指示を容易にすることである。本発明の第6の目的は、同じ画像属性を有する最大面積領域を抽出対象候補として自動抽出し、処理の簡略化と高速化を実現することである。本発明の第7の目的は、平均的輝度、色成分が所定範囲にある画像領域を領域成長核として自動抽出し、処理の簡素化と高速化を実現することである。本発明の第8の目的は、動作主体を時系列画像から抽出する際の、第1の目的に準拠した画像抽出を実現することである。

【0009】本発明の第9の目的は、動作主体を時系列画像から抽出する際の、第1の目的に準拠した画像抽出を高速に実現することである。本発明の第10の目的は、視点位置からの背景と抽出対象への平均的距離が十

分異なる場合に、第1の目的に準拠した画像抽出を実現することである。本発明の第11の目的は、対話的に大まかな指示選択により、第1の目的に準拠した画像抽出を実現することである。

#### 【0010】

【課題を解決するための手段】本発明においては、入力画像データの一次特徴を抽出する一次特徴抽出手段と、上記一次特徴を有する領域を分割する領域分割手段と、上記分割された分割領域から所定条件を満たす領域を選択する領域選択手段と、上記選択された領域を核としてその領域を成長させる領域成長手段と、上記領域成長後の画像を表示する画像表示手段とを設けている。

#### 【0011】

【作用】本発明によれば、入力画像データからまず一次的な特徴を有する領域を分割し、さらにその分割された領域のうち所定の条件を満たす領域を選択し、その選択された領域からさらにその領域を例えば二次的な特徴に基づいて拡大成長させて行き、成長後の画像が表示される。従って、表示された画像上で抽出すべき領域を大まかに位置指定するだけで、対象画像を容易に自動的に切り出し（抽出）することができる。

#### 【0012】

【発明の実施の形態】図1は本発明のシステム構成例を概念的に第1の実施の形態として示す図である。図において、1は画像入力手段で、撮像手段または画像データベースなどが用いられる。2は入力画像データ中の一次特徴データ抽出手段、3は特徴データの均質度評価手段、4は入力画像の画像記憶手段、5は特徴均質度に基づく領域分割手段、6は分割画像生成手段、7は領域成長手段、8は表示手段、9は分割領域の選択指示手段で、マウスなどのポインティング手段が用いられる。

【0013】尚、領域成長とは、画像中の所定領域内の特徴量の均質度を評価して領域の分割、併合を行うことである。本実施の形態の特徴は、入力画像中の一次特徴データを高速に抽出し、その分布に基づいて画像の領域分割を行うことにより、抽出可能性のある画像領域を概略的に抽出すると共に、次のステップとして、より精細な抽出のために、候補領域を核とする二次特徴量（一次特徴と種別、属性などが異なる特徴量）に基づく領域成長を行うことである。一次特徴データに基づいて抽出された領域には、二次特徴に基づく領域成長を行うのに必要な画像情報が含まれているものとする。

【0014】一次特徴データ抽出手段2は、後で説明する領域成長手段7による処理を行う際の核となる領域の抽出を高速に行うものであり、実際に特定対象の画像領域を抽出する際には、オペレータによる大まかな選択と指示とにより行うことができる。ここでは一次特徴データの一例として画面上各点における動きベクトルを抽出するものとして説明する。

【0015】特徴均質度評価手段3は、所定領域内の動

きベクトルの大きさと方向とに関する分散を評価値として用いる。評価する領域（ブロック）のサイズを、対象の画面内サイズに基づいて適切に設定することにより、処理時間の短縮と効率化が達成される。即ち、対象の画面内縦横サイズと比較して数分の1から10分の1程度の分割領域程度が典型的には用いられる。この値はオペレータ側で予め適宜設定するようにしてもよい。

【0016】領域分割手段5は、上記均質度の評価値が所定の閾値より小さければ、その領域を均質であるとみなし、均質度評価値が所定の許容範囲内にある領域どうしを連結する。また許容範囲外の領域は分離して一つの動きベクトルで代表することのできる領域の塊（複数可）を形成する。図7（a）に固定カメラで撮像された入力画像を示し、同図（b）に動きベクトルに基づいて行った領域分割の結果を示す。同図（c）に色成分などの二次特徴に基づく領域成長による画像切り出しの結果を示す。

【0017】尚、分割後の領域は必ずしも実際の対象物の形状を精密に反映するものでなくてもよい。従って一次特徴抽出手段2は後で用いる二次特徴抽出手段71の精度（分解能）より多少劣っていても構わないが、高速に抽出可能であることが望まれる。但し分割領域内に背景画像の一部が混在しないこと、領域成長の際、必要な画像属性（二次特徴）を全て含んでいることが望ましい。前者のためには、領域分割後、一定割合で各領域を縮小または削除したものを後続する処理に用いる分割領域としてもよい。

【0018】分割画像生成手段6は、一次特徴データに基づいて分割された複数の領域に対応してそれぞれ異なる属性の画像データ（例えば異なる色、異なるハッチングパターンなど）を付与し、表示手段8上に入力画像データと重なるように表示する。具体的には分割された領域にラベル付けを行い、同じラベル領域をマスクデータとしてマスク内を一定のパターンで塗りつぶすなどの処理を行う。このようにすることにより、オペレータは指示すべき候補領域の区別を目視により容易に行うことができ、かつ分割領域指示手段9による抽出対象の指定が確実（同じラベルで予め区分けされているから）、かつ容易となる。分割領域指示手段9としては典型的にはマウスなどが用いられるが、タッチパネル型表示装置を用いる場合は特に必要がない。

【0019】領域成長手段7は、上記のようにして一次特徴データに基づき分割され、選択された領域から二次特徴（属性）を抽出する二次特徴抽出手段71、その二次特徴に基づいて領域成長を行う場合の成長核設定手段72、各成長核に固有の二次特徴に基づいて領域成長を行う領域成長手段73、および成長後の各領域を連結する領域連結手段74から構成される。尚、領域成長とは特徴データの均質度に応じて領域の分割、併合を行うことを意味し、必ずしも一方的に併合処理を行うものでは

ない。また、二次特徴とは基本的に一次特徴と異なるものであればよく、通常は、例えば幾何学的な構造を有するもの、或いは輝度分布、色成分を処理して得られる高次の特徴を有するものがあるが、必ずしもこれらに限定されず、例えば色成分情報などが好適に用いられる。但し、一次特徴と同じ種別（ここでは動きベクトル）であっても領域成長による精細な画像分割が可能であれば問題ない（第5の実施の形態参照）。

【0020】上記のように二次特徴を抽出した後、それぞれに固有の初期領域を成長核設定手段72で設定する。色成分情報を二次特徴とする場合には、抽出された候補領域内に複数の成長すべき二次特徴が存在しうる。例えば人物の画像の場合、顔、頭髮、衣服などの部位に応じてそれぞれ平均的に異なる色相を有するが、その場合、成長核としては、一次特徴データに基づいて抽出され、かつ分割領域指示手段9で選択された候補領域の異なる二次特徴を有する輪郭部分またはその近傍領域とする。領域成長の際には予め所定サイズの近傍領域ブロック内で平均化を行ってから成長させてもよい。これによりノイズに対する耐性、画素サイズオーダーでの画像データ（属性）の変動の影響を緩和し、抽出精度を高めることができる。

【0021】図2は第2の実施の形態を示すもので、図1の画像入力手段1に撮像手段1を用いた場合の図1の他の概念的な手段2～9を具体化した手段2～9を示したものである。ここでは一次特徴データを動きベクトルとし、二次特徴データを色相とした場合である。尚、動きベクトルの検出アルゴリズムについては本発明の主眼ではないので説明を省略するが、例えば、勾配法、相関法など、「ロボットビジョン」（谷内田著、昭晃堂発行）に示されるような高速性を持つ方法が特に望まれる。

【0022】本実施の形態の適用事例としては、動作、動きを伴う対象の静止背景からの分離、または背景を含む全てが動く場合の動きベクトルの相違に基づく特定対象の背景からの分離する場合などがある。例えば撮像手段1を人体、ロボットまたは車体などに設置し、撮像手段1が動きながらあるシーンを撮像する場合には、一般的に画面全体にわたって非ゼロなる動きベクトルが存在する。このような場合には、特に背景と対象物の撮像手段1からの距離が充分離れていれば、前述のごとく動きベクトルに基づく領域分割を比較的高精度に行うことができる。特に動く対象、または周囲と異なる動きをする対象物の数が少ない場合に本実施の形態は有効である。

【0023】図3は第3の実施の形態を示すものである。本実施の形態では、画像入力手段として複眼撮像系（ステレオカメラ）を用い、一次特徴データとして左右カメラ画像間の対応点の視差分布を用いている。ステレオ画像計測法については「三次元画像計測」（井口、佐藤著、昭晃堂）などに示されている。また、本実施の形

態では撮像機構制御手段 10 を設け、複眼カメラの輻輳角、焦点距離、基線長などを視差抽出手段 2 の視差データに基づき制御することにより、抽出すべき被写体の視差が均質な領域面積が最大となるようにしている。この制御は複眼画像表示手段 8 上のラベル画像をオペレータが観測しながら行うようにしてもよい。視差は左右画像間で重複した領域についてのみ得られるので、視差値の均質な領域に分割した後の候補領域の表示は重複領域のみで行われる。領域成長は画像選択部 11 で左右画像のうち一方の抽出対象を選択して行う。また二次特徴についてはここでは色相を用いているが、これに限定されない。

【0024】図 4 は第 4 の実施の形態を示す。上記第 3 の実施の形態はステレオ画像を時系列画像データとして得ることにより、第 2 の実施の形態との併用が可能である。図 4 はそのシステム構成例を示すものである。ここでは一次特徴データ抽出手段 2 で動きベクトルと視差分布とを抽出し、それぞれの特徴データで均質に基づくクラスタリングを領域分割、ラベル付け手段 5 a、5 b で行う。図 8 (a) に入力画像を示し、同図 (b) に一次特徴（視差と動きベクトル）に基づく領域分割結果を示す。

【0025】次に偏差算出手段 12 a、12 b で分割後の領域サイズが所定値以上のクラスタについて、代表特徴量間（視差、動きベクトル）の偏差を算出し、偏差の大きい方の特徴データに基づいて領域分割され、生成されたラベル画像を抽出対象の候補領域選択の際に用いる。これは画面内の物体が複眼動画撮像手段 1 からほぼ同じ距離に分布する（視差の偏差が小）が、抽出すべき被写体が周囲に対して顕著な動きを有する場合には、動きベクトルを用いた領域分割の方が信頼度が高く、また周囲、背景との距離が複眼動画撮像手段 1 と対象との距離に比べて十分大きい（視差の偏差が大）がいずれも相対的に顕著な動きを示さない（動きベクトルの偏差が小）場合には、視差に基づく領域分割が望ましいからである。

【0026】図 8 (c) は処理結果を示し、ここでは視差値の最も大きい対象物を選択し、二次特徴微量（色成分）を用いて領域成長を行った結果を示す。これにより、対象物の動きの有無、距離分布のばらつきの大小に応じて画像抽出に適した画像データを使い分けることができる。また予め一次特徴データに基づいて使用する特徴データの種別を選択しておき、撮像モード（静止画・ステレオ撮像または動画・単眼視撮像など）を絞り込んでもよい。また一方のカメラからの動画データのみを用いる場合には第 2 の実施の形態での処理を適用する。尚、画像入力手段としては複眼撮像系に限定されず、三次元形状モデルに基づく、コンピュータグラフィックスにより生成された視差付き画像でもよい。

【0027】図 5 は第 5 の実施の形態を示す。本実施の

形態は、入力画像を固定設置された単眼カメラなどの撮像手段 1 からの時系列画像とし、一次特徴データとしてはフレーム間差分画像データを用い、一次特徴データから動きのある部分の概略領域（位置と大きさ）を抽出し、候補領域のラベル付けと表示手段 8、マウスまたはポインティング手段 9 による選択、および候補領域に対する二次特徴に基づく領域成長法を適用して画像を抽出するものである。

【0028】図 5 において、差分データ抽出手段 2 a で得られる差分画像を 2 値化手段 2 b により所定の閾値で 2 値化して得られる一次特徴データは、静止背景に対して動きベクトルの粗い分割を表し、従って均質評価を行わずにそれらを一定の基準に従って（セグメント間距離が所定画素数以下であるなど）連結処理する。典型的には所定方向に動いた物体の前方と後方に特徴データ（非ゼロデータ）が存在し、それぞれグループ化しかつそれらによって挟まれた領域を候補領域としてラベル付けし、分割領域として残す。さらに所定サイズ以下の領域を切り捨てる。二次特徴データとしては各実施の形態と同様に色相を用いているが、これに限定されるものではない。

【0029】図 6 は第 6 の実施の形態を示す。本実施の形態では、入力画像に対して多重解像・階層型の一次特徴データ（動きベクトル、視差など）抽出処理を行い、一次特徴としては低解像度データ（例えば画像を間引き処理または平均化などにより低解像化した後、動きベクトルを検出するか、または画像データをサイズの大きいブロックに分割して各ブロックごとの動きベクトルを算出するなど）、二次特徴は同じ種別の高解像度データを用いる。本実施の形態によれば、一つの特徴抽出処理アルゴリズムによる高速処理が可能である。

【0030】図 6 は一次特徴データとして動きベクトルを用いた場合のシステム構成を示す。動きベクトル均質度の評価（均質評価手段 3）から領域成長核の設定（成長核設定手段 7 2）までは低解像度で行い、領域成長（領域成長手段 7 3）は高解像度で行うことにより精細な切り出しを行う。第 2 の実施の形態のように、領域成長は均質度に基づく部分領域の併合または分離処理を意味することはいうまでもない。本実施の形態によれば、特に成長核は基本的に同じ動きベクトルで代表される領域であることから、色成分を用いた領域成長と比較して成長核設定が非常に単純化され、領域成長処理においてもより高速な処理が可能である。

【0031】

【発明の効果】以上説明したように、第 1 の発明によれば、大まかな位置を指示するだけで、任意の背景中の特定対象画像を自動的に抽出することができる。また、第 2 の発明によれば、複数の異なる種別の特徴パラメータを活用することにより、対象画像の抽出を効率的に行うことができる。また、第 3 の発明によれば、特に特徴分



Figure 1 is a block diagram of the image processing system. The components and their connections are as follows:

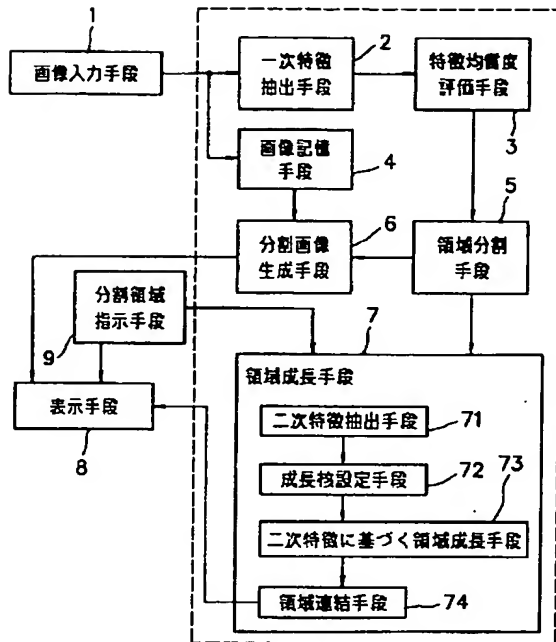
- 1** 撮像手段 (Imaging means) connects to **2a** フレーム差分抽出手段 (Frame difference extraction means) and **4** 画像記憶手段 (Image storage means).
- 2a** connects to **2b** 2値化手段 (Binarization means).
- 2b** connects to **5** 領域分割、ラベル付与手段 (Region segmentation and labeling means).
- 4** connects to **6** ラベル画像の原画上の生成手段 (Generation means of original image of labeled image).
- 5** connects to **7** 領域成長手段 (Region growth means).
- 6** connects to **7**.
- 9** マウスまたはポインティング手段 (Mouse or pointing means) connects to **7** and **8** 表示手段 (Display means).
- 7** connects to **8**.
- 7** contains the following sub-components:
  - 71** 代表色相抽出手段 (Representative color phase extraction means)
  - 72** 成長核設定手段 (Growth nucleus setting means)
  - 73** 色相に基づく領域成長手段 (Region growth means based on color phase)
  - 74** 領域連結手段 (Region connection means)

The dashed line separates the input/output section (1, 2a, 2b, 4, 5, 6, 8, 9) from the processing section (7).

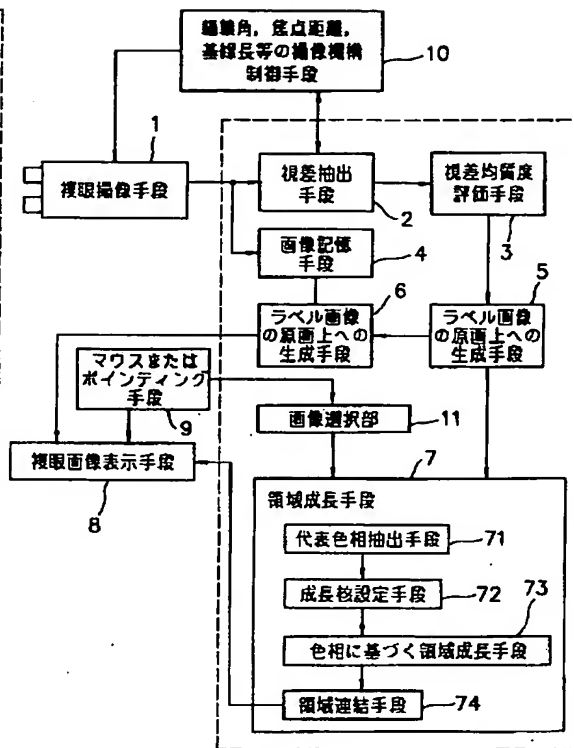


BEST AVAILABLE COPY

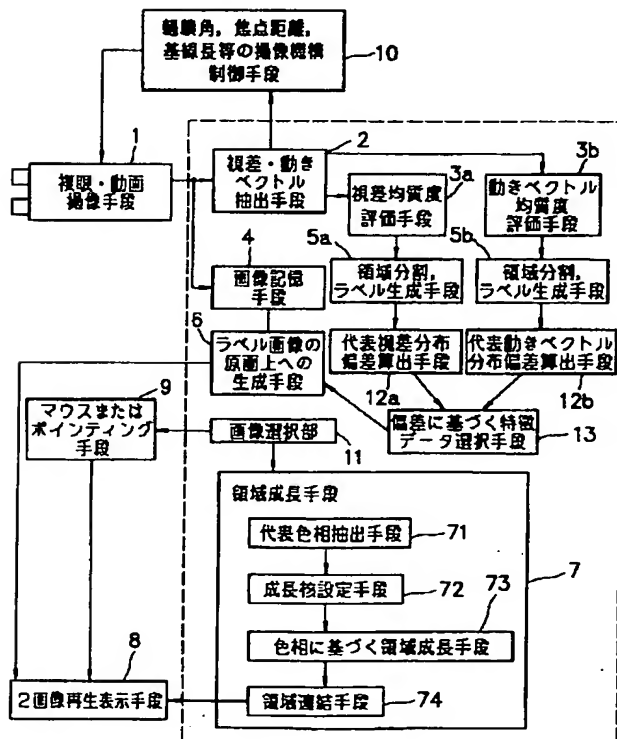
【図1】



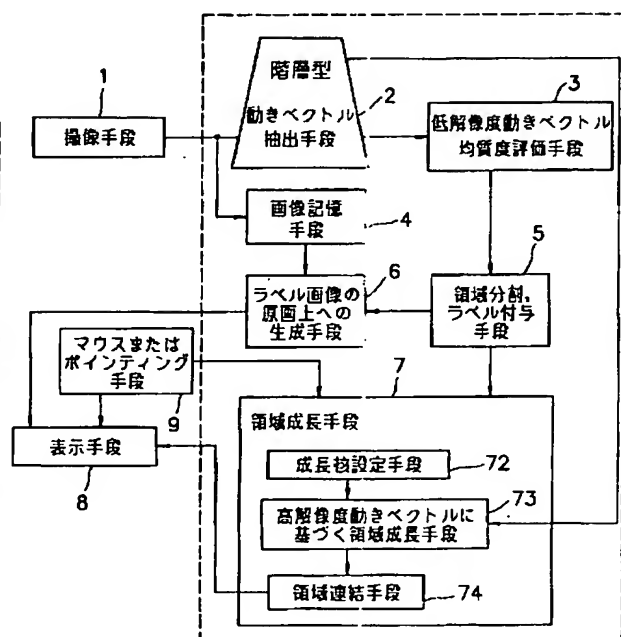
【図3】



【図4】



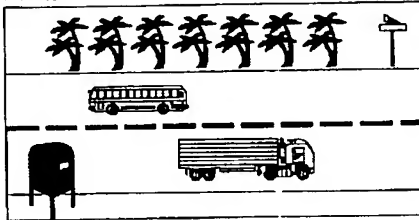
【図6】



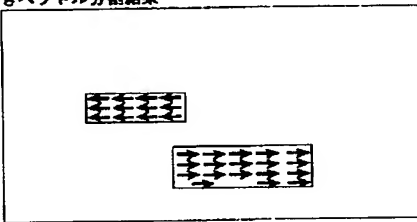
BEST AVAILABLE COPY

【図7】

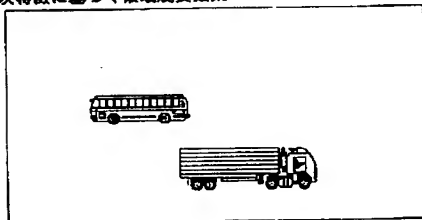
(a) 入力画像



(b) 動きベクトル分割結果

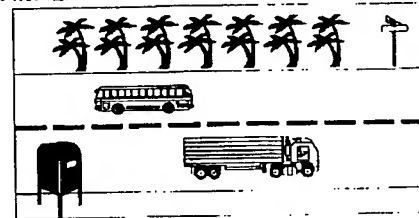


(c) 二次特徴に基づく領域成長結果

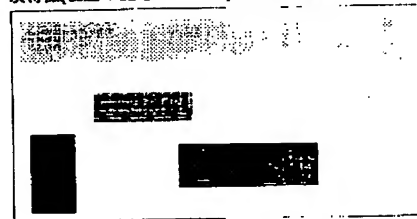


【図8】

(a) 入力画像



(b) 一次特徴(視差 + 動きベクトル)分布分割結果



(c) 二次特徴に基づく領域成長結果

